

D. Srivastava

#1

1653

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/490,291  
Input Set : A:\03665.app  
Output Set: N:\CRF3\08082000\I490291.raw

DATE: 08/08/2000  
TIME: 14:42:52

ENTERED

RECEIVED

AUG 15 2000

TECH CENTER 1600/2000

3 <110> APPLICANT: Mello, Charlene M.  
4 Arcidiacono, Steven  
6 <120> TITLE OF INVENTION: Novel Purification and Fiber Spinning Techniques for  
7 Protein Fibers  
9 <130> FILE REFERENCE: ARMY-03665  
11 <140> CURRENT APPLICATION NUMBER: 09/490,291  
12 <141> CURRENT FILING DATE: 2000-01-20  
14 <160> NUMBER OF SEQ ID NOS: 11  
16 <170> SOFTWARE: PatentIn Ver. 2.0  
18 <210> SEQ ID NO: 1  
19 <211> LENGTH: 876  
20 <212> TYPE: DNA  
21 <213> ORGANISM: Nephila clavipes  
23 <400> SEQUENCE: 1  
24 atgagaggat cgcacaccca tcaccatcac ggatccatgg ctacgcgtag aggcgggctg 60  
25 ggtggccagg gtgcagggtgc ggctgcggct gccgcggcag cggccgcagg cggtgccggc 120  
26 caaggtggct atggcggcct gggttctcag gggactagcg gttagagcgg gctgggtggc 180  
27 caggggtgcag gtgcgggtgc ggctgcggcg gcagcggcgg cagggcgggtgc cggccaaggt 240  
28 ggctatggcg gcctgggttc tcaggggact agcggtagag gcgggctggg tggccagggt 300  
29 gcagggtgcgg ctgcgggtgc cgcggcagcg gccgcaggcg gtgcgggcca aggtggctat 360  
30 ggccgctcgg gttctcaggg gactagcggg agaggcgggc tgggtggcca gggtgagggt 420  
31 gcgggtgcgg ctgcggcggc agcggccgca gccgggtgcg gccaaagggtg ctatggcggc 480  
32 ctgggttctc aggggactag cggtagaggg gggctgggtg gccagggtgc aggtgcggct 540  
33 gcgggtgcgg cggcagcggc cgcaggcggg gccggccaag gtggctatgg cggcctgggt 600  
34 tctcagggga ctacggtag aggcgggctg ggtggccagg gtgcagggtg gctgcggct 660  
35 gccgcggcag cggccgcagg cgggtccggc caaggyggct atggcggcct gggttctcag 720  
36 gggactagcg gttagagcgg gctgggtggc caggggtgcg gtgcgggtgc ggcgtccgcg 780  
37 gcagcggcgg caggcgggtg cggccaaggt ggctatggcg gcctgggttc tcaggggact 840  
38 agtgggatacc gtcgaactgc agccaagctt aattag 876  
40 <210> SEQ ID NO: 2  
41 <211> LENGTH: 291  
42 <212> TYPE: PRT  
43 <213> ORGANISM: Nephila clavipes  
45 <400> SEQUENCE: 2  
46 Met Arg Gly Ser His His His His His His Gly Ser Met Ala Ser Gly  
47 1 5 10 15  
49 Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala  
50 20 25 30  
52 Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly  
53 35 40 45  
55 Ser Gln Gly Thr Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly  
56 50 55 60  
58 Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Ala Gly Gln Gly  
59 65 70 75 80  
61 Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly Arg Gly Gly Leu  
62 85 90 95  
64 Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/490,291

DATE: 08/08/2000  
TIME: 14:42:52

Input Set : A:\03665.app  
Output Set : N:\CRF3\08082000\I490291.raw

RECEIVED

AUG 15 2000

TECH CENTER 1600/2900

```

65          100          105          110
67 Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr
68          115          120          125
70 Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala
71          130          135          140
73 Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly
74 145          150          155          160
76 Leu Gly Ser Gln Gly Thr Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly
77          165          170          175
79 Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly
80          180          185          190
82 Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly Arg Gly
83          195          200          205
85 Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala
86          210          215          220
88 Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln
89 225          230          235          240
91 Gly Thr Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala
92          245          250          255
94 Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr
95          260          265          270
97 Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly Ile Arg Arg Pro Ala Ala
98          275          280          285
100 Lys Leu Asn
101          290
104 <210> SEQ ID NO: 3
105 <211> LENGTH: 2046
106 <212> TYPE: DNA
107 <213> ORGANISM: Nephila clavipes
109 <400> SEQUENCE: 3
110 atgagaggat cgcacaccca tcaccatcac ggatccatgg ctacgcggtag aggcgggctg 60
111 ggtggccagg gtgcagggtgc ggctgcggct gccgcggcag cggccgcagg cggtgccggc 120
112 caagggtgct atggcggcct gggttctcag gggactagcg gtagaggcgg gctgggtggc 180
113 cagggtgcag gtgcggctgc ggctgccgcg gcagcggccg caggcgggtg cggccaaggt 240
114 ggctatggcg gcctgggttc tcaggggact agcggtagag gcgggctggg tggccagggt 300
115 gcagggtgcg ctgcggctgc cgcggcagcg gccgcaggcg gtgccggcca aggtggctat 360
116 ggcggcctgg gttctcaggg gactagcggt agaggcgggc tgggtggcca ggtgacaggt 420
117 gcggctgcgg ctgcccgggc agcggccgca ggcggtgccg gccagggtgg ctatggcggc 480
118 ctgggttctc aggggactag cggtcggggc gggttatggt cgggtcaaca aactagcggt 540
119 agaggcgggc tgggtggcca ggggtcagggt gcggctgcgg ctgcccgggc agcggccgca 600
120 ggcgggtccg gccagggtgg ctatggcggc ctgggttctc aggggactag cggtagaggc 660
121 gggctgggtg gccagggtgc aggtgcggct gcggctgccg cggcagcggc cgcaggcggt 720
122 gccggccaag gtggctatgg cggcctgggt tctcagggga ctacgcggtag aggcgggctg 780
123 ggtggccagg gtgcagggtgc ggctgcggct gccgcggcag cggccgcagg cggtgccggc 840
124 caagggtggt atggcggcct gggttctcag gggactagcg gtagaggcgg gctgggtggc 900
125 cagggtgcag gtgcggctgc ggctgccgcg gcagcggccg caggcgggtg cggccaaggt 960
126 ggctatggcg gcctgggttc tcaggggact agcggtcggg gcgggtatgg tccgggtcaa 1020
127 caaactagcg gtagaggcgg gctgggtggc cagggtgcag gtgcggctgc ggctgccgcg 1080
128 gcagcggccg caggcgggtg cggccaaggt ggctatggcg gctgggttc tcaggggact 1140

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/490,291

DATE: 08/08/2000

TIME: 14:42:52

Input Set : A:\03665.app

Output Set: N:\CRF3\08082000\I490291.raw

RECEIVED

AUG 15 2000

TECH CENTER 1600/2000

```

129 agcggtagag gcgggctggg tggccagggt gcagggtgcg ctgcggctgc cgcggcagcg 1200
130 gccgcaggcg gtgcgggcca aggtggctat ggcggcctgg gttctcaggg gactagcggt 1260
131 agaggcgggc tgggtggcca ggggtgcagg gcggctgcgg ctgcggcggc agcggccgca 1320
132 ggcggtgccg gccaaagtgg ctatggcggc ctgggttctc aggggactag cggtagagggc 1380
133 gggctgggtg gccagggtgc aggtgcggct gcggctgcgg cggcagcggc cgcaggcggt 1440
134 gccggccaag gtggctatgg cgccctgggt tctcagggga ctacgggtcc gggcggttat 1500
135 ggtccgggtc aacaaactag cggtagaggg gggctgggtg gccagggtgc aggtgcggct 1560
136 gcggctgccc cggcagcggc cgcagggcgg gccggccaag gtggctatgg cggcctgggt 1620
137 tctcagggga ctacgggtag aggcgggctg ggtggccagg gtgcagggtc ggctgcggct 1680
138 gccgcggcag cggccgcagg cgggtgccgg caaggtggct atggcggcct gggttctcag 1740
139 gggactagcg gtagaggcgg gctgggtggc cagggtgcag gtgcggctgc ggctgccggc 1800
140 gcagcgcccg ctagcgggtg cggccaaggt ggctatggcg gcctgggttc tcaggggact 1860
141 agcggtagag gcgggctggg tggccagggt gcagggtgcg ctgcggctgc cgcggcagcg 1920
142 gccgcaggcg gtgcgggcca aggtggctat ggcggcctgg gttctcaggg gactagcggt 1980
143 cggggcggtt atggtcgggg tcaacaaact agtgggatcc gtcgacctgc agccaagctt 2040
144 aattag
146 <210> SEQ ID NO: 4
147 <211> LENGTH: 681
148 <212> TYPE: PRT
149 <213> ORGANISM: Nephila clavipes
151 <400> SEQUENCE: 4
152 Met Arg Gly Ser His His His His His Gly Ser Met Ala Ser Gly
153 1 5 10 15
155 Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala
156 20 25 30
158 Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly
159 35 40 45
161 Ser Gln Gly Thr Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly
162 50 55 60
164 Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly
165 65 70 75 80
167 Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly Arg Gly Gly Leu
168 85 90 95
170 Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala
171 100 105 110
173 Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr
174 115 120 125
176 Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala
177 130 135 140
179 Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly
180 145 150 155 160
182 Leu Gly Ser Gln Gly Thr Ser Gly Pro Gly Gly Tyr Gly Pro Gly Gln
183 165 170 175
185 Gln Thr Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala
186 180 185 190
188 Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr
189 195 200 205
191 Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly Arg Gly Gly Leu Gly Gly
192 210 215 220

```

RECEIVED

AUG 15 2003

TECH CENTER 1600/2000

## RAW SEQUENCE LISTING

DATE: 08/08/2000

PATENT APPLICATION: US/09/490,291

TIME: 14:42:52

Input Set : A:\03665.app

Output Set: N:\CRF3\08082000\I490291.raw

```

194 Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Gly Gly
195 225                230                235                240
197 Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly
198                245                250                255
200 Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala
201                260                265                270
203 Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly
204                275                280                285
206 Ser Gln Gly Thr Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly
207                290                295                300
209 Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly
210 305                310                315                320
212 Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly Pro Gly Gly Tyr
213                325                330                335
215 Gly Pro Gly Gln Gln Thr Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly
216                340                345                350
218 Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly
219                355                360                365
221 Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly Arg Gly
222                370                375                380
224 Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala
225 385                390                395                400
227 Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln
228                405                410                415
230 Gly Thr Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala
231                420                425                430
233 Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Tyr
234                435                440                445
236 Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly Arg Gly Gly Leu Gly Gly
237                450                455                460
239 Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly
240 465                470                475                480
242 Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly
243                485                490                495
245 Pro Gly Gly Tyr Gly Pro Gly Gln Gln Thr Ser Gly Arg Gly Gly Leu
246                500                505                510
248 Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala
249                515                520                525
251 Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr
252                530                535                540
254 Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala
255 545                550                555                560
257 Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly
258                565                570                575
260 Leu Gly Ser Gln Gly Thr Ser Gly Arg Gly Gly Leu Gly Gly Gln Gly
261                580                585                590
263 Ala Gly Ala Ala Ala Ala Ala Ala Ala Ala Ala Ala Gly Gly Ala Gly
264                595                600                605
266 Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln Gly Thr Ser Gly Arg Gly

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/490,291

DATE: 08/08/2000

TIME: 14:42:52

Input Set : A:\03665.app

Output Set: N:\CRF3\08082000\I490291.raw

RECEIVED

AUG 15 2000

TECH CENTER 1600/2800

```

267      610      615      620
269 Gly Leu Gly Gly Gln Gly Ala Gly Ala Ala Ala Ala Ala Ala Ala
270 625      630      635      640
272 Ala Ala Gly Gly Ala Gly Gln Gly Gly Tyr Gly Gly Leu Gly Ser Gln
273      645      650      655
275 Gly Thr Ser Gly Pro Gly Gly Tyr Gly Pro Gly Gln Gln Thr Ser Gly
276      660      665      670
278 Ile Arg Arg Pro Ala Ala Lys Leu Asn
279      675      680
282 <210> SEQ ID NO: 5
283 <211> LENGTH: 2076
284 <212> TYPE: DNA
285 <213> ORGANISM: Nephila clavipes
287 <400> SEQUENCE: 5
288 atggctagca tgactgggtg acagcaaatg ggtcgcggtt ccatggctag cggtagaggg 60
289 gggctgggtg gccaggggtg aggtgctggc gcgctgccc cggcagcgcc cgcagggcgt 120
290 gccggccaag gtggtctatg cgccctgggt tctcagggga ctacgggtag aggcgggctg 180
291 ggtggccagg gtgcagggtg ggtgctgggt gccgcggcag cggcgcgagg cggtgcccggc 240
292 caaggtggct atggcgccct gggttctcag gggactagcg gtagagggcg gctgggtggc 300
293 caggggtcag gtgcggctgc ggtgcccgcg gcagcgcccg caggcggtgc cggccaaggt 360
294 ggctatggcg gcctgggttc tcaggggact agcggttaga gcgggctggg tggccagggt 420
295 gcaggtgcgg ctgcggctgc cgcgcgagcg gccgcaggcg gtgcccggca aggtggctat 480
296 ggcggcctgg gttctcaggg gactagcggt ccggcggtt atggtccggg tcaacaaact 540
297 agcggttaga gcgggctggg tggccagggt gcaggtgcgg ctgcggctgc cgcggcagcg 600
298 gccgcaggcg gtgcccggca aggtggctat ggcggcctgg gttctcaggg gactagcggt 660
299 agagggcgggc tgggtggcca ggggtgcagg gcggctgcgg ctgcccggcg agcggccgca 720
300 ggcgggtgccg gccaaaggtg ctatggcggc ctgggttctc aggggactag cggtagaggg 780
301 gggctgggtg gccaggggtg aggtgcgggt gcggctgcgg cggcagcggc cgcaggcggt 840
302 gccggccaag gtggtctatg cgccctgggt tctcagggga ctacgggtag aggcgggctg 900
303 ggtggccagg gtgcagggtg ggtgcgggt gccgcggcag cggccgaggg cggtgccggc 960
304 caaggtggct atggcgccct gggttctcag gggactagcg gtccggggcg ttatggtccg 1020
305 ggtcaacaaa ctacgggtag aggcgggctg ggtggccagg gtgcagggtg ggtgcgggt 1080
306 gccgcggcag cggccgaggg cggtgcgggc caaggtggct atggcgccct gggttctcag 1140
307 gggactagcg gtagagggcg gctgggtggc caggggtcag gtgcggctgc ggtgcggcg 1200
308 gcagcgcccg caggcggtgc cggccaaggt ggctatggcg gcctgggttc tcaggggact 1260
309 agcggttaga gcgggctggg tggccagggt gcaggtgcgg ctgcggctgc cgcggcagcg 1320
310 gccgcaggcg gtgcccggca aggtggctat ggcggcctgg gttctcaggg gactagcggt 1380
311 agagggcgggc tgggtggcca ggggtgcagg gcggctgcgg ctgcccggcg agcggccgca 1440
312 ggcgggtgccg gccaaaggtg ctatggcggc ctgggttctc aggggactag cggttccggc 1500
313 ggttatggtc cgggtcaaca aactagcggt agagggcggc tgggtggcca ggggtcagg 1560
314 gcggctgcgg ctgcccggcg agcggccgca ggcgggtgcc gccaaaggtg ctatggcggc 1620
315 ctgggttctc aggggactag cggtagaggg gggctgggtg gccaggggtc aggtgcgggt 1680
316 gcggctgcgg cggcagcggt cgcagggcgg gccggccaag gtggctatgg cggcctgggt 1740
317 tctcagggga ctacgggtag aggcgggctg ggtggccagg gtgcagggtc ggtgcgggt 1800
318 gcccgggcag cggccgaggg cggtgcccgg caaggtggct atggcgccct gggttctcag 1860
319 gggactagcg gtagagggcg gctgggtggc caggggtcag gtgcggctgc ggtgcggcg 1920
320 gcagcgcccg caggcggtgc cggccaaggt ggctatggcg gcctgggttc tcaggggact 1980
321 agcggtccgg gcggttatg tccgggtcaa caaactagt ggatccgaat tcgagctccg 2040
322 tcgacaagct tcgagcacca ccaccaccac cactga 2076

```

VERIFICATION SUMMARY  
PATENT APPLICATION: US/09/490,291

DATE: 08/08/2000  
TIME: 14:42:53

Input Set : A:\03665.app  
Output Set: N:\CRF3\08082000\I490291.raw

RECEIVED

AUG 15 2000

TECH CENTER 1600/2000